RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: D9 178, 188h
Source: TFW 6
Date Processed by STIC: 1-5-05

ENTERED



IFW16

RAW SEQUENCE LISTING

DATE: 01/05/2005

PATENT APPLICATION: US/09/478,188H

TIME: 11:02:32

Input Set : A:\407T-896010US new.txt
Output Set: N:\CRF4\01052005\I478188H.raw

3 <110> APPLICANT: SHEN, BEN LIU, WEN CHRISTENSON, STEVEN D. 5 STANDAGE, SCOTT 8 <120> TITLE OF INVENTION: GENE CLUSTER FOR PRODUCTION OF THE ENEDIYNE ANTITUMOR 9 ANTIBIOTIC C-1027 (ps.6) 11 <130> FILE REFERENCE: 407T-896010US 13 <140> CURRENT APPLICATION NUMBER: 09/478,188H 14 <141> CURRENT FILING DATE: 2000-01-05 16 <150> PRIOR APPLICATION NUMBER: 60/115,434 17 <151> PRIOR FILING DATE: 1999-01-06 19 <160> NUMBER OF SEQ ID NOS: 119 21 <170> SOFTWARE: PatentIn Ver. 3.3 23 <210> SEQ ID NO: 1 24 <211> LENGTH: 42000 25 <212> TYPE: DNA 26 <213> ORGANISM: Artificial Sequence 28 <220> FEATURE: 29 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic C-1027 gene cluster sequence 32 <220> FEATURE: 33 <223> OTHER INFORMATION: orf; relative position 658-11 35 <220> FEATURE: 36 <223> OTHER INFORMATION: orf; relative position 1478-930 38 <220> FEATURE: 39 <223> OTHER INFORMATION: orf; relative position 2713-1649 41 <220> FEATURE: 42 <223> OTHER INFORMATION: orf; relative position 3238-2851 44 <220> FEATURE: 45 <223> OTHER INFORMATION: orf; relative position 4971-3442 47 <220> FEATURE: 48 <223> OTHER INFORMATION: orf; relative position 5982-7478 50 <220> FEATURE: 51 <223> OTHER INFORMATION: orf; relative position 9900-7573 53 <220> FEATURE: 54 <223> OTHER INFORMATION: orf; relative position 11349-9982 56 <220> FEATURE:

57 <223> OTHER INFORMATION: orf; relative position 28590-29588

60 <223> OTHER INFORMATION: orf; relative position 29632-31197

63 <223> OTHER INFORMATION: orf; relative position 31280-32590

59 <220> FEATURE:

62 <220> FEATURE:

65 <220> FEATURE:

Input Set : A:\407T-896010US new.txt
Output Set: N:\CRF4\01052005\I478188H.raw

66 <223> OTHER INFORMATION: orf; relative position 32809-34392 68 <220> FEATURE: 69 <223> OTHER INFORMATION: orf; relative position 35274-34458 71 <220> FEATURE: 72 <223> OTHER INFORMATION: orf; relative position 17924-16653 74 <220> FEATURE: 75 <223> OTHER INFORMATION: orf; relative position 16653-15919 77 <220> FEATURE: 78 <223> OTHER INFORMATION: orf; relative position 15922-14690 80 <220> FEATURE: 81 <223> OTHER INFORMATION: orf; relative position 14643-14212 83 <220> FEATURE: 84 <223> OTHER INFORMATION: orf; relative position 13012-14079 86 <220> FEATURE: 87 <223> OTHER INFORMATION: orf; relative position 12835-11351 89 <220> FEATURE: 90 <223> OTHER INFORMATION: orf; relative position 25564-24986 92 <220> FEATURE: 93 <223> OTHER INFORMATION: orf; relative position 24702-23566 95 <220> FEATURE: 96 <223> OTHER INFORMATION: orf; relative position 22878-21424 98 <220> FEATURE: 99 <223> OTHER INFORMATION: orf; relative position 21407-19926 101 <220> FEATURE: 102 <223> OTHER INFORMATION: orf; relative position 19929-19267 104 <220> FEATURE: 105 <223> OTHER INFORMATION: orf; relative position 19191-18031 107 <220> FEATURE: 108 <223> OTHER INFORMATION: orf; relative position 35938-35516 110 <220> FEATURE: 111 <223> OTHER INFORMATION: orf; relative position 27214-28593 113 <220> FEATURE: 114 <223> OTHER INFORMATION: orf; relative position 25815-27170 116 <220> FEATURE: 117 <223> OTHER INFORMATION: orf; relative position 23546-22875 119 <220> FEATURE: 120 <223> OTHER INFORMATION: orf; relative position 35274-34458 122 <220> FEATURE: 123 <223> OTHER INFORMATION: orf; relative position 37559-38938 125 <220> FEATURE: 126 <223> OTHER INFORMATION: orf; relative position 40986-39367 128 <400> SEQUENCE: 1 129 gtcgactcta gaggatcccg ggtgcggagt aggggttacg gacgaaggag gggtgcccgg 60 130 cgacgcctgc ggcgaagggc ggttccttga gttcgaggcc ggtggcgagg acgacgtggt 120 131 ccgcgtcgag gatctgcgtg tcggggagcg gcccagggcg cagccctcg gtcaggtacg 180 132 gggtgaggcc cctgacggtc acctcgaagc agcggtcgtg ggaccgggcg tcgagcgcct 240 133 eccegteege ttecacaagg acgaegeegg gacaggaete eegtgeggee tegaceagte 300 134 gggcgtcgag gtagtcctgg aagatgcggc ggggggcggg gccctgttcg gtgaacttcc 360 135 acgaageeca gegeegggge cagtegegee ggteggeete etggttggee cagttgatga 420

Input Set : A:\407T-896010US new.txt
Output Set: N:\CRF4\01052005\I478188H.raw

136 agtcgagcac gtcctcgcgg aacaccgaca tcctgccggc ctggatattg aagacgtggt 480 137 cccaggggtt gccgtcacgg tgataggcga cgccggccga gcggtaggcg gcgccgct 540 138 ccaggaggac gacttccagc ggtcttctcg cgaaatgaag caggcgtatc gcggtcgccg 600 139 tgcctgccag gcccgccct acgaccagca ccctggggcg cgcacccgtc atgcccatga 660 140 agecteccee getgacteag ggeggeget egegggetee egteggtgte etegetgact 720 141 ggaagtteee tgaeetggeg teaacteeae tgateegtaa ggggategeg ggagtggata 780 142 cgggtcaggt cgtgcacgat cgtggcacca gacagatcac cacgtcgata ggcactcgtg 840 143 agecgegece ggggetegae ggggegggg aceggeaggg geggeegegt gateageegg 900 144 agcctgtccg ggggcgtgcg tgcggggcgt cagctgtcga tgtcgggaac gccagggacg 960 145 tegatetegg tgegggegta gtggttgaag tagttggtgt agaggtteae ggeeaegtgg 1020 146 acgaagacct cggcgagctc ggtgtccgtc catccctgtg ccacggccgc gttccacgag 1080 147 gegteagacg cetegeceae ttegeeggeg atetecetgg ceaectggae cagtgetteg 1140 148 agetteaegt egtegeeggg egteeeeegg egaategeea eggteteete eagegtgaaa 1200 149 cccgcgacct tcgccgacac cgtgtgcgcc gcctggcagt acgcgcacgc gtcgaccgcg 1260 150 cccacggcga gggcgatcgc ctcgcgtgtg cgggcgtcga acgttccatg ttcggcgacg 1320 151 gctccggtga tcgcggcgta ggtttccagg accacggggg aatgggccat tcccccgtgg 1380 152 atgttgagca ctcgcccgaa ccgcttctcc agtcggcgca ggatgtctcc gccggctgcg 1440 153 ggtgcggtgt cgatggtgtg gacgggaatc cgcggcatgg gaatgcctct cctcgtagtg 1500 .154 atgggagttc ctcgtccctc cagtctgccc aagcacctcc cccggtgagc tgtcccggcc 1560 155 gccctccggc cccttctagg caggtcgccc ggtggtgcgg ccccaggacg tcacctcgcc 1620 156 gcaccaccgg gagccccgag gggcgaggtc agaggccgag cacctcctcg gccagggcgg 1680 157 tgccccgaac acgggcctcg atcttggcga aggccaggtc gcgtgtggtg gaggtgtcgt 1740 158 cggcgaacgg ggagaagccg cagtcgtcgc aggttcccag ttgctcgacg gggatgtagc 1800 159 gggcggcgag caggatgcgg tcgcgtacct gctcgggggt ctcgaccact gggtcgatcg 1860 160 ggtcggtcac cccgaggaag acgcgggcgg cagggggcag gtggtcacgg acgatgctca 1920 161 ggacccgctc ggggtccgct tcgccggcca gttcgagata gaagttgccc gccttgagct 1980 162 ggaagagett gggeageagt teggegtagt egatgtegag getgtgegtg gagteetggt 2040 163 cgccgccggg gcaggtgtgt acgccgatgc gggcggtttc ctcggcgctg aagcgcccca 2100 164 ggacttegtt gttgagggeg atgaagtegt egaggaegee geegetgggg tegagettga 2160 165 gggacagceg ceceteggtg aagtegaget ggaccaegtg tgeeceegeg tecaggcage 2220 166 ctcggatgtc ggcttcggcc tcgtcggcga ggtcgcgcag gaactgctcg cgggggtagc 2280 167 cctcgatggg agtggcgggg tagaggaggc tgagggcgga gggtgcgatg accgcctgct 2340 168 tcagggggcg gtccgtgagc tgccgtgcgg cgcgcagata ggtttcggcc cgcacctggt 2400 169 ageggaaggg ceettgggtg atgetgggga getgeegggt gtgeeegtet gegaagggga 2460 170 tgacagcgcc gtcgggcgag agggtgtcga ggccggtcac ggggtaggtg gcgaagctcg 2520 171 gcttggactg ttcaccgtcc acgaggacgg ggctgccgac tcgttccagt cgtgtcaggg 2580 172 tgtccgcgac ggcctgttcc tgctgtttgg ccaggtccgt ggcgtccagg gttccctggg 2640 173 catgegegge aagggegtge aggagtgteg eggagegegg aaggetgeeg ateggeteag 2700 174 tggcgatggt catggccgaa gagtagggaa gaggctgggt ttcgaaccac cgcaaagctt 2760 175 tgattgccgc tttttcaggg gaagttgatg cgaagtcgcc gagcggcgga acgtgctgat 2820 176 gtatgggggg cgggaggagc ctgcggggtt ctaggagccg gtcgcggcca cggtggagga 2880 177 ggtgcccagc tgggagcggg gggtcttttc gccgacgcgg ttgggctcga tggtgcgggg 2940 178 gtcgacggcc tetecggggg cacettgccg gtagacgcet teggggtegg agtcceggte 3000 179 atgggggagc aggaagaaga cccggcgccg gtacagaccg ctgtccgggt ccgcttcggc 3060 180 gtcggccccg agttcgatgt agccgatcat gcggccgtcg cgggcgtagc gcggcttgtt 3120 181 cttgcgccgg ggggtcttgt ccagggcctg gcggacgtag tcgagtccct cgggatcttc 3180 182 gagccacacg accttcgcct cgtgaacgag atcgctgtcg gtcagtagcg agctcatggc 3240 183 ggcgacetet cettegtegg egtgeacegg gtggggaage ggtgeetgeg tgatgtgtgt 3300 184 tegtetgegg eggtgggeeg eagtggtgeg gaeegeeegt ggtgeeggtt eteggeeaaa 3360

Input Set : A:\407T-896010US new.txt
Output Set: N:\CRF4\01052005\I478188H.raw

185 gcacgggcag gtacgtcctg gggcactcac atcgtagatg gggtccgctt ccgcagggca 3420 186 gtgcctccgg tcggaggacg ttcattcgtc ggctgccaga gcgaggttgg ggtagaactt 3480 187 ccggccgttg gatttgatca tgtcggcagg tgaggcgagg cccacttcct ggcggacccg 3540 188 ggtggcgaag gcacgggcgg tcccggggcg gatgccttca ctgtgtgcgc accaggtgct 3600 189 gtaggaegtg tagagaagge cetgttegae gegtageteg etgttetegg ggtegtggag 3660 190 gcagcactcg gcgaggaagc ggccgatgtg gtcctcggtg ttcgcgtatg cgctggtggc 3720 191 gatgeggace eggtegggge eggegagtgt gtegegggtg gegaggtage ggegggeece 3780 192 tteggtgage cagtgeagga teeeggggee etegteetgg acgagttega cagecaggtt 3840 193 gtcgatcttg cgttcgtcgg ggacgatccg ttcgaagggc aggaggcgga tgcggcgca 3900 194 gaaggegaag cegeeggtgg agaceteggg geggtggttg ceeageagee acagettgtg 3960 195 cgtgggtgtg aaggagaaat agtcctgccg catgcggcgg gccttgatct tgtcaccgcc 4020 196 ggtcagcagg cggacgcgcg cctcgtcgaa gcggtcgttg ggcttgagct cgctgcacac 4080 197 gatgaggcgg cggccgtgga gttcggtgag ctcggtggag tgttcggagt atgcgccacg 4140 198 gtccatgagg aaacccggcg gggctgcgtc ggcgtagtcg ccgagaatct ggatcatcac 4200 199 gtcgaggaga acggatttgc cgttctttcc ctggccgtgg agaaagggca gcacctgcgc 4260 200 cccgacgtca ccggtgatgg agtagccgag aaggaggtgg aggaagtcga tcatctcccg 4320 201 cccttcggcg tcactgccga aggtgtcttc gaggaaacgg tgccagcggg gggtggggat 4380 202 gtcctggggg gaggcgctgg tggcgcggga gtggaagtcc cgggtggggt cgggcttgcg 4440 203 catacggccg ttgcggaggt cgaccactcc gtcaggggtg cacagggcgt aggggtctcc 4500 204 gtcgagggtg tcgggatcga gggagaggtc gggagaggcc tttgcctggg tgaggagcgc 4560 205 cttcataccg gtcgtcgaca gggtgcggcg tttgtggtgg tgcagttccc ggtcggtgaa 4620 206 cagecegegg ggategetge egggeatete eteegeeate teteeggeag eccaeaggge 4680 207 agettteteg ceteeggeee getteeaceg gtageegtee caggagtace ageceaggee 4740 208 ctccacgtgc cggaactggt cacggtagag acggacgaag agcttggcgt tgccgcggtc 4800 209 ggtcaggctg gcgggaatet cgcccgcctc ccaggcggtc gcggcgacgg gggcctcggg 4860 210 ageggeetgg acagggagga geggegetgg ggeeggggtg gtttegaggg ecageatetg 4920 211 ctgagcggcg gcagttgcgt caaagcgagg gccctcggcg ctgctgctca tggacgtcct 4980 212 tegagatgga geggteggge ggteceeget gegggaaegg catgaatgat etteeeggtg 5040 213 cggacagagt gccaggggca gcgcatgtgc ggggggacaa cggcccgttt cggacgaggg 5100 214 ccggccgacg gggggaagca ggggccggca accgggtggc ggggcggcgt gagcgagggc 5160 215 acgageggee eggtaegggg ggaagggete gteteteegt ggggeggeac gttgtggtee 5220 216 tegteegtea gettgegtet ggetteagee teetgaceee caataaggeg aaagetgetg 5280 217 gtcaagcatc tttcgtgaca ctcggcgagg gactgaaggg actgtctttc ggaatgagtg 5340 218 tagggggttg tegggtgggg accgegeete gaeteeeegg eggaegggat etgtteggte 5400 219 ggtcccttgg gtccctcccc ggatcgcggc agggacccaa gggggcggtg cggcgggcgg 5460 220 tcggtgaggg gccccggtgg agggactgag ggtctgtatg gagcgataag agggtctgaa 5520 221 ggggcggaga gagtttcggt ccctgcgttg agtccctggt catcaccgca ggtcagaggg 5580 222 gttttgaggg gtgaaaaagg gactgaaggg actcaacttc cccattatga gctgagtaga 5640 223 agaaagcagt atgacgatat cggcgcctac atacgcgcgc gtacatagtg agcttataat 5700 224 geggaagttg agteeettea gteeetttte gtggggtegt ateeeetetg aetgegttga 5760 225 ccgtcgccgc tccgcgcagg gaccgaagag ggaccaagtc cctgcgcggg gcgggcgacg 5820 226 gtaatcgtgc agtgccccct cccccgtttc ccacagcgag tcgtcgctcc cctgtgaggc 5880 227 cggagagggt cctagaaccc ctcaggggcc gttctgtggc cctctgggcc tcctcctggc 5940 228 catttacccc atgggggcgc ttgggggcgt caggagggct tgtgagggct ctgccgggaa 6000 229 gtggcggatt gcgcatggca ggagatgccc cgacagcggc cgggaatcga cgatgtcccc 6060 230 cgacccctat ccagcgtccg ctgatcctca ggaggcagac cttgcaggct ccagaagcga 6120 231 agaacggccg gtccccggag cagccgcagg aagagcggat cgtcctggac gtatggctgg 6180 232 egaactacce gttccccacc tatgacggge gtgacttcct cgctccgctg cgcgagcggg 6240 233 cggcggagtt cgagcgcgcc cacccccgat accgggtcga catcaacggc cacgacttct 6300

Input Set : A:\407T-896010US new.txt
Output Set: N:\CRF4\01052005\I478188H.raw

234 ggaccatece egagaaggtg gegegeeca eegeggaggg caggeeteeg cacatagegg 6360 235 gctactacgc caccgacagc cagttggcgc gggacgcgcg caggcccgac gggaagccgg 6420 236 tetteacete ggtggaggee gegttggeeg geeggaegga gataetggga caeceggtgg 6480 237 tggtggagga cetegaceee gtggtgegeg aeteetaete gtteggggge gagttggtgt 6540 238 egetgeeget caeggteace accatgetet getaegeeaa etecteeete etegegegeg 6600 239 ccggtgttcc ggagttgccc cgtacctggg atgaggtcga agcagcctgc caggcggtgg 6660 240 ccagcgtega cggggggccc ggtcacggaa tcacctgggc caacgacggc tgggttttcc 6720 241 agcaggccgt cgcccttcag aacggggtgc tgaccgatca ggacaacggc cgctccggct 6780 242 ccgccacgac ggtggacgtc acatcggacg agatgctgga ctgggtccgc tggtggacgc 6840 243 acctccatga gcgcggccat tacctctaca cgggcgggcc ctcggactgg ggcggggcgt 6900 244 tegaggettt egteeageag aaggtegeat teacettega etegteeaag geegeeeggg 6960 245 aactcatcca ggccggtgca caggccggtt tcgaggtcgc ggtgttcccg ttgcccagga 7020 246 acqcqaaqqc cccqqtaqcq qqccaqcccq tctcqqqaqa ctccctqtqq ctqqccqcqq 7080 247 gactegaega gaccaegeag gaegggetge tegeteteae eeagtaeetg ateageeegg 7140 248 ccaacgccgc ggactggcac cgcaccaacg gtttcgtacc ggtgaccggc gcggccgggg 7200 249 aactgctgga agcgacaggc tggttcgacc gccggccgca gcaacgggtg gccggggagc 7260 250 agttgaagge gteegaeegg teaeeggegg egeteggege getgetegge gaettegegg 7320 251 ccgtcaacga ggtcatcacc gcagcgatgg acgatgtcct gcgcagtgga gcggaccccg 7380 252 cgaaggeett egeegaagee ggegtggeeg eecageaact getegatgee tacaaegeee 7440 253 ggaaccgete eggateeggg acceeeteeg eegtetgaga teeggtaeeg gggeaeaggg 7500 254 gegeegeege eegettteee ggegggeae tggeeggggg acatgetete eegeeeeegg 7560 255 caggacgtag ggtcaacceg cetgegeett caggtggegg egeagataet caeeggteag 7620 256 ggaggaatee geggegagea ggteettegg tgtgeeggtg aagaegatet egeegeeete 7680 257 ccgtcccccg tcgggaccca ggtcgatgat ccagtcggcc tgctgcacca catcgaggtt 7740 258 gtgctcgatg accacgacgg tgttcccggc ctcgacgagc ccgtccagga gcttcagcag 7800 259 ggtgtcaacg tccgacatgt gcagcccggt ggtgggctcg tccaggacat agaccgtgcc 7860 260 egtgeggtge agetggtegg caagtttgat eegetgeagt teaeegeegg agaggetgga 7920 261 aageggetgg cecaggetga ggtacecaag acegaegteg aegagagege geagtttegg 7980 262 cagcagggcc ttctcggtga agaactcgac ggcctcgtcg gcgggcagct ccaggacgtc 8040 263 egegategae ttecegegaa getggtgete eaggaeeteg ggettgaage ggegeeete 8100 264 acagacaccg cagtgcgtgg tcaccggatc catgaaggcc agctcggtga tgatgacccc 8160 265 geggeeetgg eacteetege aegaeeeett ggagttgaag etgaacageg aggegttege 8220 266 gccggtctcc ttcgcgaaca gcttgcgcag cgggtccatc aggccgaggt aggagaccgg 8280 267 tgtggagege gaegaggegg egategegga etggtegaea aagaeegegt eggggtgege 8340 268 etecatgaat geeceggaga teaggetget ettgeeggaa eeegeeaeee eggteaeege 8400 269 ggtcagcaca ccggtgggca cggccacgga gacctgcttc aggttgtgga gatccgcgtt 8460 270 ctccacggtc ageteeceg tgggegggeg gaceteetee ttcacgeggg eccecegeeg 8520 271 cagageetee ceggteeggg tettegeett eegeagette gegaaggace cetegaacae 8580 272 gatetegece cegtgeacte cegeceeggg acegacateg aegatgtggt eggegatete 8640 273 gatcacatcg gggtcgtgct cgacgaccag cacggtgttc cccttgtcgc gcagcgcgcg 8700 274 cagcaggteg ttgageegee ceaegtegeg egggtgeagg eegatgetgg getegtegaa 8760 275 gatgtacgtg agcccggcca gaccactgcc gaggtggcgc accatcttca gccgctgccc 8820 276 ctcgccccc gagaggtcgg ccgtgggcct gtccagggtc aggtagccga gcccgatgga 8880 277 cacgateege teeagggeeg tgegegege tttegegaga ggggeagegg ceggeteegt 8940 278 gacgccggcg agcacctccg tgaggtcgcg gacctccatg ctcgagtagt cggcgatgtt 9000 279 cttgccgtcg atccggacgt cgagcgcggc ggcgttgagc cgcgcgcccc ggcaggaggg 9060 280 acagacteeg teggtgaega aaegttegat gaeetegege ttgeggtege teagegeget 9120 281 gaggtegege ttgaggttga geegetegaa eeggteggee aaceeetegt agttegtetg 9180 282 gaacteggtg etettggtet teagegteae ettecegeeg gtgeegegea geagegtgte 9240.

Input Set : A:\407T-896010US new.txt
Output Set: N:\CRF4\01052005\I478188H.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:104; N Pos. 9 Seq#:105; N Pos. 13

Invalid <213> Response:

Use of "Artificial" only as "<213> Organism" response is incomplete, per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:119

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/478,188H TIME: 11:02:33

DATE: 01/05/2005

Input Set : A:\407T-896010US new.txt

Output Set: N:\CRF4\01052005\1478188H.raw

L:2517 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:104 after pos.:0 L:2558 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:105 after pos.:0